PACIFIC JOURNAL OF MATHEMATICS
Vol. 43, No. 2, 1972

ALGEBRAS OF NORMAL MATRICES

GEORGE MAXWELL

A classical theorem of matrix theory asserts that a commuting set of complex normal matrices can be simultaneously unitarily diagonalised. In this paper, this result is generalised, both for the field of complex numbers and for more general fields. Namely, a commuting set of normal matrices is replaced by a subalgebra composed entirely of normal matrices. The structure of such subalgebras is determined and results on simultaneous diagonalisation are deduced. In the complex case, these subalgebras turn out to be commutative. However, even in the real case there are noncommutative examples.

1. Normal subalgebras. Let F be a field with an involution J, V a finite dimensional vector space over F and ϕ a left hermitian form on V such that

$$\phi(x, x) = 0 \quad \text{implies } x = 0 . \tag{1}$$

In particular, ϕ is nondegenerate so that every endomorphism T of V has a unique adjoint w. r. t. ϕ, defined by the equation

$$\phi(Tx, y) = \phi(x, T^*y) . \tag{2}$$

We call a subalgebra A of $\mathrm{End}_F(V)$ normal if it satisfies

$$\begin{aligned} &\text{(a)} \quad T \in A \quad \text{implies} \quad T^* \in A \\ &\text{(b)} \quad T^*T = TT^* \quad \text{for all} \quad T \in A . \end{aligned} \tag{3}$$

Our first aim is to determine the structure of such normal subalgebras.

The purpose of assuming (1) is to obtain the property

$$T^*T = 0 \quad \text{implies} \quad T = 0 . \tag{4}$$

Indeed, if $T^*T = 0$, we have $\phi(Tx, Tx) = \phi(x, T^*Tx) = 0$ so that $Tx = 0$ for all $x \in V$. From properties 3(a) and (4), a well known argument [6] leads to the fact that A has no nil ideals. In our context, this means that A must be semisimple. Furthermore, if B is a minimal ideal of A, so is B^*, and thus either $B^* = B$ or $B^*B = 0$, but the latter possibility is precluded by (4). It is therefore sufficient to determine the structure of a simple normal subalgebra.

PROPOSITION 1. *Suppose R is a ring with unit element $1 \neq 0$ and $*$ is an involution of a matrix ring $M_n(R)$ with the property $XX^* =$*

X^*X for all $X \in M_n(R)$. Then either (i) $n = 1$ *or* (ii) $n = 2$, *R is commutative and $*$ is the involution*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^* = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{5}$$

Proof. Linearing the identity $XX^* = X^*X$, we obtain

$$XY^* + YX^* = X^*Y + Y^*X;$$

replacing Y by Y^*, this can be written as

$$[X, Y]^* = -[X, Y]. \tag{6}$$

Let $E_{ij}(r)$ be the matrix with r in the (i, j)th position and zeros elsewhere. Suppose $n \geqq 3$; if $i \neq j$, we can write $E_{ij}(r) = E_{ik}(1)E_{kj}(r) = [E_{ik}(1), E_{kj}(r)]$ for some $k \neq i, j$. Therefore $E_{ij}(r)^* = -E_{ij}(r)$ by (6); but then $E_{ij}(r)^* = E_{kj}(r)^*E_{ik}(1)^* = E_{kj}(r)E_{ik}(1) = 0$, an absurdity.

If $n = 2$, we can write $E_{12}(r) = [E_{11}(1), E_{12}(r)]$ so that $E_{12}(r)^* = -E_{12}(r)$. Since $E_{11}(r) = E_{12}(1)E_{21}(r)$, we have $E_{11}(r)^* = E_{22}(r)$; the involution is thus given by (5). Furthermore, writing $E_{11}(rs) = E_{11}(r)E_{11}(s)$ and applying $*$, we obtain $E_{22}(rs) = E_{22}(sr)$ so that $rs = sr$ and R must be commutative.

PROPOSITION 2. *Suppose D is a division ring, finite dimensional over its center Z and $*$ is an involution of D such that $dd^* = d^*d$ for all $d \in D$. Then either $D = Z$ or D is a quaternion algebra over Z and $*$ is the standard involution.*

Proof. Let K be the subfield of Z left fixed by $*$ and L some algebraic closure of K. The extended involution $(d \otimes \alpha)^* = d^* \otimes \alpha$ on $D \otimes_K L$ has the same property as $*$.

If $K = Z$, $D \otimes_K L$ is isomorphic to $M_p(L)$ for some integer p. By Proposition 1, $p \leqq 2$ so that D is either Z or a quaternion algebra over Z (see, e.g., [1. p. 146]). If $K \neq Z$, we have $Z \otimes_K Z \cong Z \oplus Z$, so that $D \otimes_Z L \cong D \otimes_Z (Z \otimes_K Z) \otimes_Z L \cong D \otimes_Z L \oplus D \otimes_Z L \cong M_p(L) \oplus M_p(L)$ for some integer p. If $*$ induces an involution on each of the factors $M_p(L)$, we again have $p \leqq 2$. However, if $p = 2$, we see from (5) that $*$ must leave central elements fixed, which is not true for $D \otimes_K L$. Therefore $p = 1$, i.e. $D = Z$. If $*$ interchanges the two factors $M_p(L)$, then each is forced to be commutative so that once again $p = 1$.

It remains to verify that in case D is a quaternion algebra over Z and $K = Z$, $*$ can only be the standard involution. If char$(Z) \neq 2$, D has a basis $\{1, i, j, ij\}$ such that $i^2 = \alpha$, $j^2 = \beta$ and $ij = -ji$ for some $\alpha, \beta \in Z$. Since $2\beta i = [ij, j]$ and $2\alpha j = [i, ij]$, (6) implies that $i^* = -i, j^* = -j$ so that $*$ must be the standard involution. If char $(Z) = 2$, the relations are instead $i^2 = \alpha, j^2 = j + \beta$ and $ij = ji + i$ for some $\alpha, \beta \in Z$. Since $i = [i, j]$ and $ij = [j, ij]$ we have $i^* = i$ and $(ij)^* = ij$; but $\alpha j = i(ij)$ so that $\alpha j^* = (ij)^* i^* = iji = \alpha j + \alpha$ i.e. $j^* = j + 1$, showing that $*$ is again the standard involution.

The preceding proofs could have been somewhat shortened by appealing to a recent result of Amitsur [3], which says that a semiprime ring with an involution $*$ satisfying a polynomial identity $p(X_1, \cdots X_n, X_1^*, \cdots, X_n^*) = 0$ of degree d satisfies a "standard identity" of degree $2d$. In our case, the polynomial identity is $X_1^* X_1 - X_1 X_1^* = 0$, of degree 2, so that the standard identity is of degree 4. Now a well-known result of Kaplansky [7] implies that if the ring is also primitive, it is at most 4-dimensional over its center. However, we would still have to determine, as above, the possibilities for $*$, the knowledge of which is important in the sequel.

PROPOSITION 3.
(a) *If J is non-trivial, a simple normal subalgebra A is a finite field extension of F; its involution $*$ extends J.*
(b) *If J is trivial, A can also be a quaternion division algebra over a finite field extension of F, in which case $*$ must be the standard involution.*

Proof. Suppose A is isomorphic to $M_n(D)$, where D is some finite dimensional division algebra over F. By Proposition 1, either (i) $n = 1$ or (ii) $n = 2$, D is a field and $*$ corresponds to the involution (5). However, the latter violates (4) since, for example, $E_{11}(1)^* E_{11}(1) = 0$; therefore A is a division algebra. By Proposition 2, A is either a field or a quaternion algebra over its center. Furthermore, in the latter case $*$ must be the standard involution, which is certainly trivial on F, so that J itself had to be trivial.

Turning to the classical cases, let us suppose that F is either $\boldsymbol{R}$ or $\boldsymbol{C}$ and ϕ is the standard hermitian form on $V = F^n$.

COROLLARY 1, *In the complex case, a normal subalgebra is isomorphic to a product of copies of $\boldsymbol{C}$, each with the standard involution.*

COROLLARY 2. *In the real case, a normal subalgebra is isomorphic to a product of copies of $\boldsymbol{R}$, $\boldsymbol{C}$ and $\boldsymbol{H}$, the latter two occurring with the standard involution.*

Proof. It is only necessary to explain why a factor consisting of $\boldsymbol{C}$ with the trivial involution could not occur in the real case. This is a consequence of a property stronger than (4):

$$(7) \qquad \sum T_i^* T_i = 0 \quad \text{implies that all} \quad T_i = 0\,,$$

enjoyed by $*$ but violated by such a factor. Indeed, if $\sum T_i^* T_i = 0$, we have $\phi(\sum T_i^* T_i x, x) = \sum \phi(T_i x, T_i x) = 0$ for all $x \in V$; since all summands are non-negative, we must have $\phi(T_i x, T_i x) = 0$ and hence $T_i = 0$.

2. Simultaneous diagonalisation. Let A be a normal subalgebra of $\mathrm{End}_F(V)$ and consider V as a left A-module. One sees at once from (2) that if W is a submodule of V, so is $W^\perp$; in view of (1), we have $V = W \oplus W^\perp$. Induction now shows that V is the orthogonal sum of simple submodules, which are isomorphic to simple factors of A.

Using Corollaries 1 and 2 of Proposition 3, we can immediately obtain diagonalisation results in the classical situations.

PROPOSITION 4. *In the complex case, there exists an orthonormal basis of V w.r.t. which the matrices of all elements of A are diagonal.*

PROPOSITION 5. *In the real case, there exists a partition* $\dim V = n_1 + 2n_2 + 4n_3$ *and an orthonormal basis of V w.r.t. which the matrices of all elements of A consist of n_1 diagonal elements, followed by n_2 blocks of the form*

$$(8) \qquad \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

and n_3 blocks of the form

$$(9) \qquad \begin{pmatrix} \alpha & -\beta & -\gamma & -\delta \\ \beta & \alpha & -\delta & \gamma \\ \gamma & \delta & \alpha & -\beta \\ \delta & -\gamma & \beta & \alpha \end{pmatrix}.$$

Proof. If a simple A-submodule is isomorphic to $\boldsymbol{C}$, it has a basis of the form $\{x, i\cdot x\}$, which is orthogonal since $\phi(x, i\cdot x) = \phi(i^*\cdot x, x) = -\phi(i\cdot x, x) = -\phi(x, i\cdot x)$. We may suppose that $\phi(x, x) = 1$, but then $\phi(i\cdot x, i\cdot x) = \phi(x, i^* i\cdot x) = \phi(x, x) = 1$, so that the basis is orthonormal. The action of $\boldsymbol{C}$ on such a basis is given by blocks of the form (8). Similarly, if an A-submodule is isomorphic to $\boldsymbol{H}$, it has a basis of the form $\{x, i\cdot x, j\cdot x, ij\cdot x\}$, which can once again be assumed orthonormal and yields blocks of the form (9).

Such diagonalisation results are usually stated for a commuting set $\{T_i\}$ of normal endomorphisms rather than for a normal subalgebra. To deduce them from our results, we first enlarge the set $\{T_i\}$ to $\{T_i, T_i^*\}$, which is still commuting in view of the following well-known result [9]:

PROPOSITION 6. *In the real or complex case, if a normal endomorphism T commutes with an endomorphism S, it also commutes with S^*.*

Secondly, we form the commutative subalgebra generated by $\{T_i, T_i^*\}$, which is clearly normal, and apply propositions 4 and 5. In the non-classical situations, the results of §1 still enable us to produce diagonalisation theorems, although these can of necessity be more complicated. We shall confine ourselves to some remarks about the case when $F = \boldsymbol{Q}$ and ϕ is the standard hermitian form on $V = \boldsymbol{Q}^n$.

PROPOSITION 7. *The possible factors of a normal subalgebra must be of the following types*:
(a) *a totally real finite extension $K/\boldsymbol{Q}$, with the trivial involution.*
(b) *an extension $K(\sqrt{-\alpha})/\boldsymbol{Q}$, where K is as in* (a) *and α is totally positive, with the involution $\sqrt{-\alpha} \to -\sqrt{-\alpha}$.*
(c) *a quaternion algebra $(-\alpha, -\beta)$ over K, where K is as in* (a) *and α, β are totally positive, with the standard involution.*

Proof. Let A be a simple factor. We go back to proposition 3. If * induces the trivial involution on A, every $T \in A$ is hermitian and therefore has totally real eigenvalues-hence A is of type (a). When * is not trivial, the fixed subfield K of * is of type (a) by the same argument. If $K \to \boldsymbol{R}$ is some imbedding, then, regarding $\boldsymbol{R}$ as a K-algebra, one proves as before that the involution $(a \otimes \lambda)^* = a^* \otimes \lambda$ on the extended algebra $A \otimes_K \boldsymbol{R}$ enjoys property (7). Therefore the images of α or α and β must be positive in $\boldsymbol{R}$.

The problem of determining which totally real extensions $K/\boldsymbol{Q}$ can actually occur as factors of type (a), say, has been studied by Bender [4] and seems quite difficult. For example, $\boldsymbol{Q}(\sqrt{d})/\boldsymbol{Q}$ occurs if and only if d is a sum of 2 squares in $\boldsymbol{Q}$.

3. The infinite dimensional case. In this paragraph, we shall prove that in some infinite dimensional situations normal subalgebras are necessarily commutative.

Firstly, suppose that H is a complex Hilbert space and $B(H)$ is the algebra of bounded operators on H. The analogue of Proposition 6 for elements of $B(H)$ has been proved by Fuglede [5] and later generalised by Putnam [10] to

PROPOSITION 8. *If S and T are normal operators and R is an operator such that $TR = RS$, then $T^*R = RS^*$.*

One can use this result to prove

PROPOSITION 9. *A normal subalgebra A of $B(H)$ such that A^2 is dense in A (for example if $1 \in A$) must be commutative.*

Proof. Suppose $S, T \in A$; since $(ST^*)S = S(T^*S)$, Proposition 8 implies that $(ST^*)^*S = S(T^*S)^*$ or $T(S^*S) = (S^*S)T$ (this idea occurs in Kaplansky [8]). Now replace S by $S + R^*$, with $R \in A$. After subtraction, one concludes that T commutes with $(RS)^* + RS$ i.e. with all the hermitian elements of A^2. Since A^2 is dense in A and every element of A can be written in the form $S + iT$ where S and T are hermitian elements of A, we conclude that T commutes with every element of A.

Secondly, we return to an arbitrary field F and consider an arbitrary F-algebra Ω with an involution $*$, satisfying $(\alpha . x)^* = \alpha^J . x^*$. Let $b(\Omega)$ be the quotient of $\Omega \otimes_k \Omega$ by the subspace generated by all elements of the form $ab \otimes c - a \otimes bc$ and $ba \otimes c - a \otimes cb$. The obvious map $\beta_\Omega \colon \Omega \times \Omega \to b(\Omega)$ is called the universal bitrace on Ω. It may happen that $b(\Omega)$ is not isomorphic to K, for example if $\Omega^2 = 0$. Since Ω has an involution, it is actually more convenient to work with a "twisted" version of the bitrace: $\langle a, b\rangle = \beta_\Omega(a^*, b)$. This is a left sesquilinear (w.r.t. J) map on Ω, universal w.r.t. the properties $\langle ab, c\rangle = \langle b, a^*c\rangle$ and $\langle ba, c\rangle = \langle b, ca^*\rangle$. By analogy with [2], Ω may be termed an H^*-algebra if

$$\langle a, a\rangle = 0 \quad \text{implies} \quad a = 0 \, . \tag{10}$$

For such algebras, the analogue of Proposition 8 can be proved purely formally from the identity

$$\begin{aligned} &\langle c^*a - bc^*, c^*a - bc^*\rangle - \langle ac - cb, ac - cb\rangle \\ &\quad = \langle aa^* - a^*a, cc^*\rangle - \langle bb^* - b^*b, c^*c\rangle \, , \end{aligned} \tag{11}$$

a special case of which goes back to von Neumann [11]. For its proof, note first that $\langle ab, cd\rangle = \langle bd^*, a^*c\rangle = \langle c^*a, db^*\rangle$. Then

$$\begin{aligned}&\langle c^*a - bc^*, c^*a - bc^*\rangle \\ &\quad = \langle c^*a, c^*a\rangle - \langle bc^*, c^*a\rangle - \langle c^*a, bc^*\rangle + \langle bc^*, bc^*\rangle \\ &\quad = \langle aa^*, cc^*\rangle - \langle cb, ac\rangle - \langle ac, cb\rangle + \langle c^*c, b^*b\rangle .\end{aligned}$$

Similarly, $\langle ac - cb, ac - cb\rangle = \langle a^*a, cc^*\rangle - \langle ac, cb\rangle - \langle cb, ac\rangle + \langle c^*c, bb^*\rangle$. Subtraction yields (11).

PROPOSITION 10. *If J is nontrivial, a normal subalgebra A of an H^*-algebra Ω such that $A^2 = A$ must be commutative.*

Proof. One can use the same argument used in the proof of Proposition 9, with the following remark. Since J is nontrivial, there exists $\theta \in F$ such that $\theta^J \neq \theta$; then every $x \in A$ can be written in the form $x_1 + \theta\cdot x_2$, where $x_1 = (\theta\cdot x^* - \theta^J\cdot x)/(\theta - \theta^J)$ and $x_2 = (x - x^*)/(\theta - \theta^J)$ are hermitian elements of A.

In conclusion, we add a remark regarding the property

$$aa^* = a^*a,\ bb^* = b^*b,\ ac = cb \quad \text{implies} \quad c^*a = ab^* \tag{12}$$

in arbitrary rings with involution. Two of its special cases are

$$aa^* = a^*a,\ ac = ca \quad \text{implies } c^*a = ac^* \tag{13}$$

and

$$aa^* = a^*a,\ ac = 0 \quad \text{implies} \quad c^*a = 0 . \tag{14}$$

However, one can get an example in which both (13) and (14) hold but (12) does not, by taking $K = \boldsymbol{Q}$, $\alpha = 2$ in

PROPOSITION 11. *Let K be a field of characteristic $\neq 2$, α a nonzero element of K and $*$ the involution*

$$\begin{pmatrix} a & b \\ c & d\end{pmatrix}^* = \begin{pmatrix} a & -\alpha c \\ -b/\alpha & d\end{pmatrix}$$

of $M_2(K)$. Then (i) (13) *is true in $M_2(K)$,* (ii) (14) *is true iff α is not a square and* (iii) (12) *is true iff α is not a sum of* 2 *squares.*

We omit the full proof, but give the counterexample for (12): suppose $\alpha = \beta^2 + \gamma^2$ and let

$$X = \begin{pmatrix} 1 & \beta \\ \beta/\alpha & 1\end{pmatrix},\ Y = \begin{pmatrix} 2 & -\gamma \\ \gamma/\alpha & 0\end{pmatrix},\ Z = \begin{pmatrix} \alpha & -\alpha\gamma \\ \beta & 0\end{pmatrix} .$$

Then $XX^* = X^*X$, $YY^* = Y^*Y$, $XZ = ZY$ but $X^*Z \neq ZY^*$.

REFERENCES

1. A. A. Albert, *Structure of algebras*, Amer. Math. Soc. Colloq. Publ., no. 24, 1961.
2. W. Ambrose, *Structure theorems for a special class of Banach algebras*, Trans, Amer. Math. Soc., **57** (1945), 364-386.
3. S. A. Amitsur, *Identities in rings with involution, Israel* J. Math., **7** (1969), 63-68.
4. E. A. Bender, *Characteristic polynomials of symmetric matrices*, Pacific J. Math., **25** (1968), 433-441.
5. B. Fuglede, *A commutativity theorem for normal operators*, Proc. Nat. Acad. Sci., U.S.A. **36** (1950), 125-134.
6. I. Kaplansky, *Rings of operators*, Benjamin, New York, 1968.
7. ———, *Rings with a polynomial identity*, Bull. Amer. Math. Soc., **54** (1948), 575-580.
8. ———, *Products of normal operators*, Duke Math. J., **20** (1953), 257-260.
9. M. Marcus and H. Minc, *A survey of matrix algebra and matrix inequalities*, Allyn & Bacon, Boston, 1964.
10. C. R. Putnam, *On normal operators in Hilbert space*, Amer. J. Math., **73** (1951), 357-362.
11. J. Von Neumann, *Approximative properties of matrices of high finite order*, Portug. Math., **3** (1942) 1-62.

Received June 17, 1971 and in revised form June 6, 1972.

UNIVERSITY OF BRITISH COLUMBIA